基于 MATLAB 的一维光谱可视化与分析工具的设计与实现*

张珊瑚, 邱丹, 聂嘉潞, 李崇晟, 王习东, 田海俊, 刘高潮，郑胜

（三峡大学天文与空间科学研究中心 湖北 宜昌 443002）

摘要: 随着 LAMOST 光谱巡天项目的不断推进，国际上积累的光谱数量已经步入千万量级。因科学研究需求或大量低信噪比光谱的不断产出，在光谱处理过程中，常涉及到大量人工识谱的任务。为方便用户高效便捷地进行人工识谱，本文基于 MATLAB 编程技术设计并实现了一套一维光谱可视化与分析工具。该工具包括谱线标识、滤波去噪、红移测量、等值宽度估算等功能。该工具操作简单、直观，在特殊天体搜寻，LAMOST“UNKONWN”光谱的人工处理等方面具有较高的使用价值。

关键字：MATLAB；一维光谱；工具；红移测量；等值宽度估算

Design and Implementation of
Visualization and Analysis Tool for 1-D Spectrum Based on MATLAB

Zhang Shanhu, Qiu Dan, Nie Jialu, Li Chongsheng, Wang Xidong, Tian haijun,
Liu Gaochao, Zheng Sheng

(Center for Astronomy and Space Science, Three Gorges University, Yichang, Hubei 443002)

Abstract: With the development of LAMOST spectroscopic Sky Survey project, the accumulated spectra in the world have stepped into tens of millions of magnitudes. Due to the need of scientific research or the continuous output of a large number of low signal-to-noise ratio spectra, a large number of manual spectral identification tasks are often involved in spectral processing. In order to facilitate users to carry out manual spectrum recognition efficiently and conveniently, this paper designs and implements a set of one-dimensional spectral visualization and analysis tools based on MATLAB programming technology. The tool includes spectral line identification, filtering and denoising, redshift measurement, and equivalent width estimation. The tool is simple, intuitive operation. It has high application value in the search of special celestial objects and the artificial processing of LAMOST "UNKONWN" spectra.

Keywords: MATLAB; 1-D spectra;tools; redshift measurement; equivalent width estimation

1 引言

随着各个大型光谱巡天项目不断地开展，最近十几年来，天文学领域已收集了丰富的天体光谱，这些光谱为科学研究提供了重要的数据基础。我国自主研发的大天区面积多目标光纤光谱天文望远镜(Large Sky Area Multi-Object Fiber Spectroscopy Telescope，LAMOST)[1]自 2012 年正式巡天到 2019 年 3 月，已经获取了 1125 余万条光谱，其中“UNKNOWN”光谱 100 余万条。随着巡天观测的深入，“UNKONWN”光谱会越来越多。

这些“UNKONWN”光谱大多数因为信噪比较低，传统的算法无法处理；也可能有极少数奇异的天体，由于谱型太奇怪，传统的算法无法识别，如毛晓燕[2]提出了基于加权滤波的低信噪比 LAMOST 光纤光谱信号降噪。由于每条光谱的获取成本较高，且这些光谱本身可能含有丰富的科学价值，不能舍弃它们，因此如何识别挖掘这些“UNKONWN”光谱的

*基金项目：国家自然基金（11873034，U1731124）
收稿日期： 年-月-日；修回日期：年-月-日
作者简介：张珊瑚，女，本科生。研究方向：数学与应用数学。
通讯作者：田海俊，男，博士。研究方向：天文信息学，星系天文学。Email：hjtian@lamost.org

科学价值是我们急需解决的问题。一些天文学家可以对这些低信噪比光谱进行分类和参数值估计，但是如此巨量的数据仅靠少数的专家利用有限的时间和精力完成人工处理是不可能的，必须开发自动化的光谱分析软件。“工欲善其事，必先利其器”，只有借助良好的光谱分析软件，天文学家才能便捷地对光谱进行分析和研究，从而提高工作效率和科学产出。

国内外专家设计了诸多一维光谱数据可视化与分析软件。比如，欧洲空间局虚拟天文台研发团队开发了 VOSpec[3]，可以对光谱进行查询、分析、拟合光谱能量分布；美国空间望远镜科学研究所开发的 SpecView[4]，可以对光谱进行查询、可视化与分析；德国虚拟天文台团队与捷克共和国科学院的天文研究所共同开发的 SPLAT[5]，可以显示、比较、修改和分析天文光谱数据，并且能查询和下载光谱；法国天体物理学和行星学研究开发的 CASSIS [6]，可以查询光谱、估算谱线的柱密度等；中国科学院国家天文台开发的 ASERA [7]，可以识别类星体光谱和测量红移；雷国洪等人设计了一套基于数据挖掘的专家识谱平台[8]，可以对光谱数据进行管理、可视化与分析和科学研究与科学普及等。这些软件确实能挖掘天体光谱的科学价值，然而我们发现 VOSpec、SpecView、SPLAT 对 LAMOST 光谱数据存在不兼容现象；ASERA 主要针对类星体的光谱而设计，且不能进行滤波去噪、等值宽度估算等。

为此，本文基于 MATLAB[9]设计并实现一套一维光谱的可视化与分析工具，重点完成 LAMOST “UNKNOWN” 的人工处理，其中包括光谱的人工分类、光谱的红移和等值宽度等特征物理参量的测量、特殊天体搜寻等科学目标。设计的可视化界面力求简单直观，适合普通学生使用，通过发动学生力量，逐渐消化不断产生的“UNKNOWN”光谱数据。

2 一维光谱可视化与分析工具的总体结构设计

本文将设计一个基于 MATLAB 开源的一维光谱可视化与分析工具，利用 MATLAB 设计的图形用户界面（MATLAB GUI）实现一维光谱的可视化，并结合相关算法实现对光谱进行平滑去噪、红移测量、等值宽度估算等分析功能。一维光谱可视化与分析工具的总体结构设计如图 1 所示：

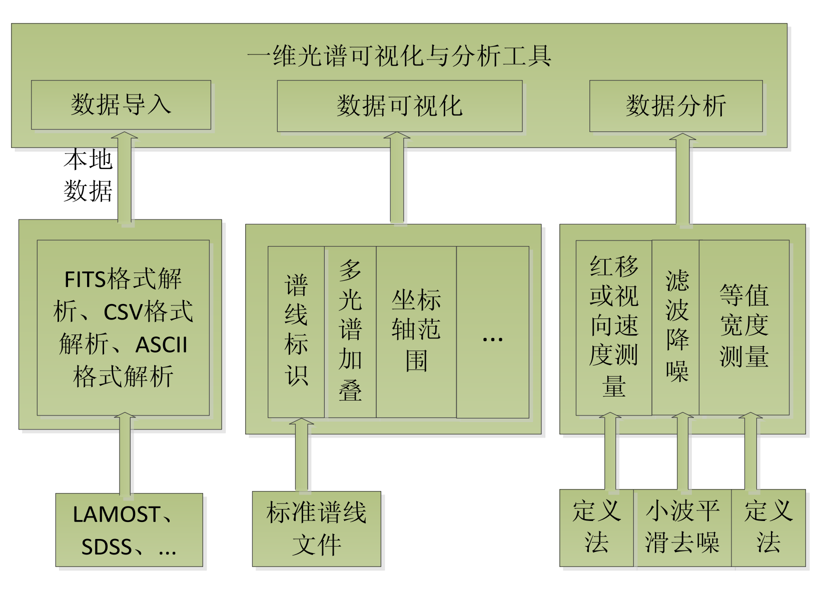


图 1 一维光谱可视化与分析工具总体结构设计图

Fig. 1 Overall architecture design of 1-D spectral visualization and analysis tool

一维光谱可视化与分析工具分为三个部分，分别为数据导入，数据可视化和数据分析。

（1）数据导入：用户将本地光谱数据导入到一维光谱可视化与分析工具中，光谱文件格式主要分为 FITS 格式、CSV 格式和 ASCII 格式。LAMOST，SDSS 等大型巡天项目的光谱文件通常为 FITS 格式。

（2）数据可视化：导入光谱数据后，通过 MATLAB 程序设计实现光谱数据可视化。光谱数据可视化包括特征谱线标识、多光谱加叠、光谱横纵坐标轴设置、谱线类型设置等。标准谱线文件是在静止坐标系下的特征谱线的波长列表，通过导入标准谱线文件来进行谱线标识。

（3）数据分析：数据分析包括红移或视向速度的测量、滤波降噪处理以及等值宽度的估算等。运用定义法计算红移或视向速度，运用小波平滑去噪算法进行光谱去噪处理，运用定义法测量光谱的等值宽度。

3 一维光谱可视化与分析工具的具体实现

可视化界面设计分为桌面设置、光谱数据导入设置、数据可视化与分析设置三部分，数据分析功能主要包括光谱去噪处理、光谱红移或视向速度的测量和等值宽度测量。这些功能都将通过 Matlab 编程技术来实现。

3.1 桌面设置

3.1.1 桌面颜色设置

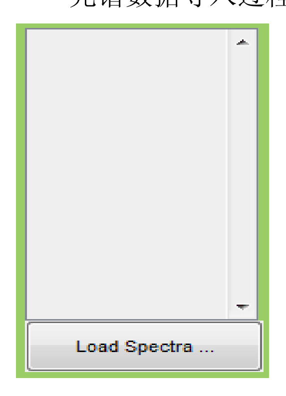
用户可根据个人的喜好选择不同的颜色界面，系统设计了 6 种颜色可供选择,分别为绿色（Summer Greens）、蓝色（Winter Blues）、橙色（Earthy Copper）、粉色（Subtle Pinks）、灰色（Gray Tones）和黑色（Dark Blacks）选项。

3.1.2 Logo 图标设置

在图形用户界面的左上角是系统的 Logo 展示区，用户可设置自己的 Logo 图标。

3.2 数据导入设置

光谱数据导入过程如图 2 所示：



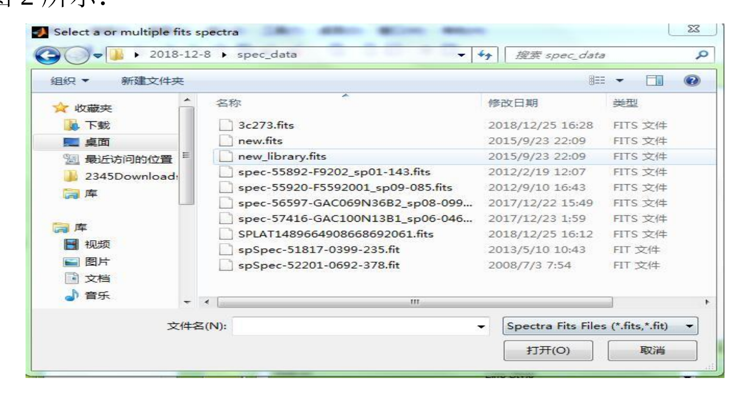


图 2 FITS 光谱文件导入过程图

Fig. 2 Diagram of importing FITS spectral file

系统仅支持导入一维光谱数据，且主要支持 FITS 文件格式，用户需要将本地 FITS 光谱数据上传至系统中。

3.3 数据可视化设置

3.3.1 谱线标识

谱线标识能辅助用户识别光谱上主要的特征谱线，在静止坐标系下测得的常见的特征谱线波长如表 1 所示：

表 1 常见特征谱线的波长列表（单位：λ/nm）

Table 1 The wavelength list of some common spectral lines (unit: λ/nm)

特殊谱线	O VI	Ly	N V	O I	C II	SiIV	CIV	HeII	OIII
静止坐标	1033.8	1215.7	1240.8	1305.5	1335.3	1397.6	1549.5	1640.4	1665.8
特殊谱线	AlIII	CIII	NeIV	MgII	Ne V	NeVI	OII	He I	S II
静止坐标	1857.4	1908.7	2439.5	2799.1	3346.8	3426.8	3727.1	3889	6718.3
特殊谱线	N I	N II	Li	K	H	G	MgI	NaI	CaII'
静止坐标	6529	6549.9	6707.9	3934.8	3969.6	4305.6	5176.7	5895.6	8500.4

点击 Speclines 按钮，系统将根据用户设置的红移值，将特征谱线以虚线的形式显示出来。

3.3.2 多光谱加叠

点击 Hold on 按钮，图形用户界面可以同时导入多条一维光谱，并对这些光谱进行比较分析，点击 clear 按钮清除可视化界面上的光谱图像。多光谱加叠显示如图 3 所示：

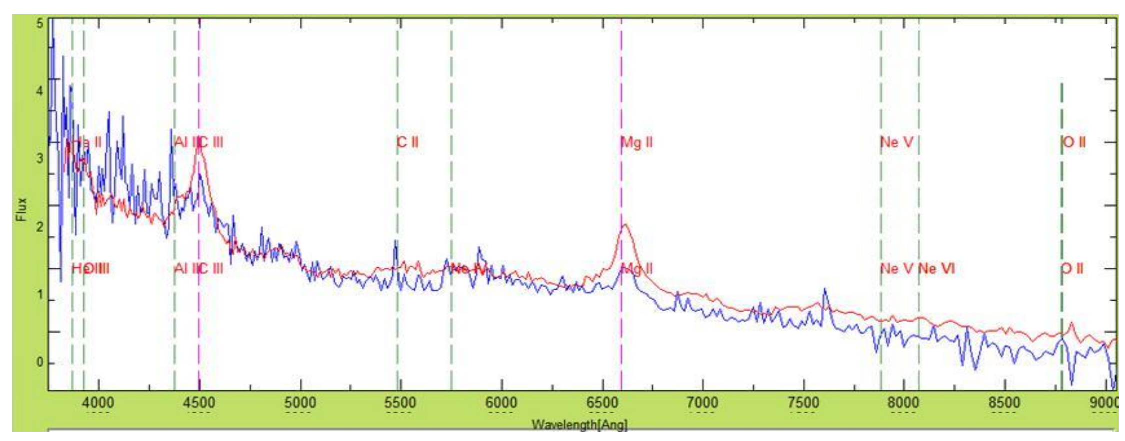


图 3 多光谱加叠图

Fig.3 Diagram of multi-spectrums overlapping

图 3 中可视化界面显示了两条不同的类星体光谱图像，用红色虚线显示的特征谱线有较强的吸收线，特征谱线 MgII 发生红移后波长坐标在 6500~6750nm 之间，特征谱线 CIII 发生红移后波长坐标在 4500nm 左右。

数据可视化还包括对光谱图像进行网格划分、改变光谱横纵坐标轴的范围、设置谱线类型和设置图形用户界面的标题，用户可以根据自己的需求在可视化界面的编辑框进行设置。

4 光谱数据分析功能的实现

根据彭广民[10]等人的小波去噪方法，李天超[11]等人的等值宽度算法以及刘蓉[12-13]等人的红移测量和谱线证认方法，本文运用 MATLAB 编程技术实现了光谱平滑去噪处理、红移测量和等值宽度测量等光谱分析功能。

4.1 光谱平滑去噪算法

小波去噪方法既能去除图像中的噪音，又可以尽可能地保留图像局部的细节，因此小波

去噪法可以很好的对光谱进行平滑去噪处理。

4.1.1 光谱平滑去噪法的原理

设 (x,y) 为图像中点的坐标， $f(x,y)$ 为含有噪声的图像， $e(x,y)$ 为图像上噪音干扰，σ 是噪声方差。图像去噪后可以表示为：

$$g(x,y)=f(x,y)-\sigma e(x,y),(x=3700,3701,...,9000;y=-10,-9,...,10)$$

小波去噪的主要步骤：

（1）对 $f(x,y)$ 作小波变换，得到各尺度的小波系数 $w_{(j,k)}$；

（2）设定各分解层阈值，对小波系数 $w_{(j,k)}$ 进行阈值处理，得到小波估计系数 $\hat{w}_{(j,k)}$；

（3）重构小波系数，再利用小波估计系数 $\hat{w}_{(j,k)}$ 进行重新构造得到去噪处理后的图像 $\hat{f}(x,y)$。

4.1.2 阈值函数选择

阈值的确定和对小波系数的阈值处理是小波去噪的关键，这里选择半软阈值法进行阈值处理：

$$\hat{w}_{(j,k)}=\begin{cases}\operatorname{sgn}\left(w_{(j,k)}\right)\left(\left|w_{(j,k)}\right|-\alpha\lambda\right) & (w_{(j,k)}\geq\lambda)\\ 0 & (w_{(j,k)}<\lambda)\end{cases}$$

其中，$w_{(j,k)}$ 为处理前的小波系数，$\hat{w}_{(j,k)}$ 为处理后的小波系数，λ 为阈值，且 $0<\alpha<1$。

4.1.3 光谱平滑去噪结果

本文将光谱平滑程度设置成 6 个等级，分别为 3、5、7、9、11 和 13（单位：pix），平滑程度为 3 pix 和 11 pix 的光谱图去噪效果如图 4 所示：

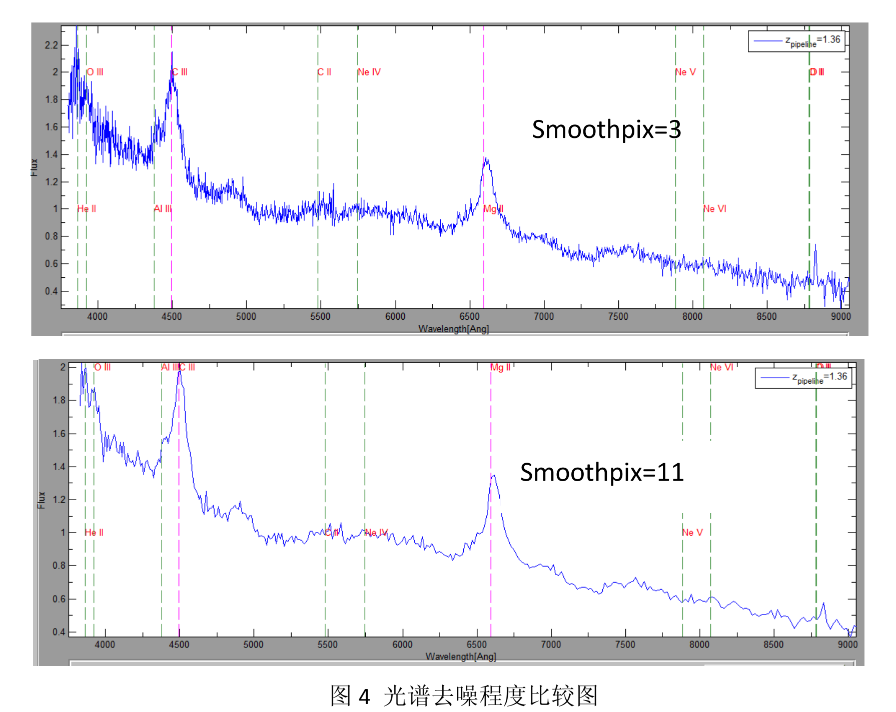


图 4 光谱去噪程度比较图

Fig. 4 Comparisons of spectral denoising

从上图可以看出，平滑程度为 11pix 的光谱图比平滑程度为 3pix 的光谱图更平滑，视觉

效果更好，但是 11pix 可能会损伤部分微小的光谱特征，比如 AlIII 等。因此，用户需要根据自己的需求来选择平滑程度。

4.2 红移或视向速度测量算法

在物理学和天文学领域，红移现象指物体的电磁辐射由于某种原因使波长被拉伸的现象，具体表现为光谱的谱线朝红端移动了一段距离，即波长变长，频率降低。红移现象目前多用于天体的移动以及规律的预测上，红移值是河外天体最重要的物理参数之一。

4.2.1 红移测量法的原理

设谱线的静止波长为λ'，观测波长为λ， z 为红移值，则

$$\lambda = (1+z)\lambda'$$

$$z = \frac{\lambda}{\lambda'} - 1$$

谱线的静止波长λ'如表 1 所示，根据可视化界面的特殊谱线位置确定观测波长λ后，将观测波长输入到编辑框，点击 enter 键，得到特殊谱线的红移值。

4.2.2 红移测量结果

本文以三条光谱测量为例，运用一维光谱可视化与分析工具测得光谱的特征谱线 CIII 和 MgII 的红移值如表 2 所示:

表 2 光谱特征谱线红移值

Table 2 of Spectral redshift

光谱文件名	C III	MgII
spSpec-51817-0399-235.fit	1.3566	1.3675
spSpec-52201-0692-378.fit	1.3570	1.3656
spSpec-55920-F5592001-sp09-085.fits	1.3587	1.3636

4.3 等值宽度测量算法

等值宽度，即与吸收（或发射）谱线轮廓和连续谱之间所包围的面积相当的高度为 1 的矩形的宽度。

4.3.1 等值宽度测量法的原理

下面以一观测光谱上 CIII 谱线为例，概述天体特征谱线等值宽度的测量方法（如图 5 所示）。

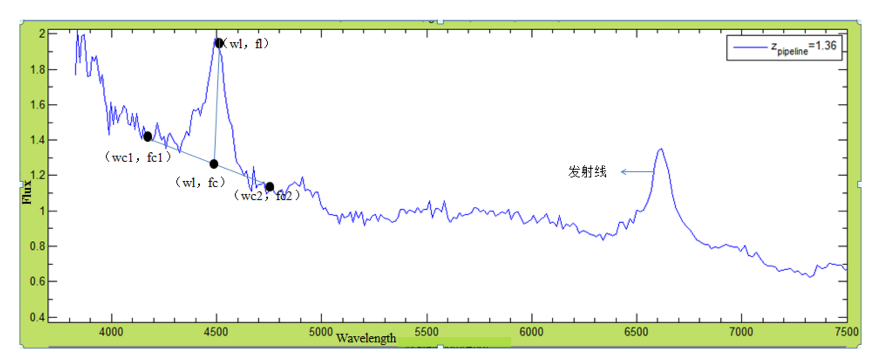


图 5 CIII 的等值宽度测量示意图

Fig. 5 Chart of equivalent width measurement for CIII

已知特征谱线峰值的取值范围是$\left[\left(wl_1, fl_1\right),\left(wl_n, fl_n\right)\right]$，峰值左边点的取值范围是$\left[\left(a_1,b_1\right),\left(a_j,b_j\right)\right]$，峰值右边点的取值范围是$\left[\left(c_1,d_1\right),\left(c_k,d_k\right)\right]$，峰值左边点的坐标$\left(wc1, fc1\right)$为：

$$wc1=\frac{\sum a_j}{j}$$

$$fc1=\frac{\sum b_j}{j}$$

峰值右边点的坐标$\left(wc2, fc2\right)$为：

$$wc2=\frac{\sum c_k}{k}$$

$$fc2=\frac{\sum d_k}{k}$$

$\left(wl_i, fc_i\right)$是坐标$\left(wc1, fc1\right)$和坐标$\left(wc2, fc2\right)$连接线上的一点，$\left(wl_i, fc_i\right)$的值为：

$$fc_i=\frac{\left(fc2-fc1\right)}{\left(wc2-wc1\right)}\left(wl_i-wc1\right)+fc1$$

等值宽度的计算结果为：

$$ew=\int_{wl_1}^{wl_n}\left(1-\frac{fc_i}{fl_i}\right)dwl$$

4.3.2 等值宽度测量结果

本文以三条光谱测量为例，运用一维光谱可视化与分析工具测得光谱的特征谱线 CIII 和 MgII 的等值宽度值如表 3 所示:

表 3 光谱特征谱线等值宽度

Table 3 of Spectral Equivalent Width

光谱文件名	C III	MgII
spSpec-51817-0399-235.fit	58.592	73.492
spSpec-52201-0692-378.fit	53.797	67.268
spSpec-55920-F5592001-sp09-085.fits	59.297	35.366

5 总结与展望

LAMOST、2dF、SDSS 等国际上大型巡天项目不断获取大量的光谱数据，这些光谱为天文学研究提供了重要的数据基础。为方便用户深度挖掘这些光谱的科学价值，本文基于 Matlab 研发了一套一维光谱可视化与分析工具。该工具具备以下几个方面的优势：

（1）针对一维光谱数据，集成多种可视化与分析算法，方便用户简单快捷地计算出光谱的红移或视向速度、等值宽度等重要物理参数。

（2）界面简洁，操作步骤简单，用户只需要掌握可视化界面的操作方法，就可以对光谱进行分析处理。

（3）能像 SPLAT-VO 等项目一样，实现一维光谱数据可视化与分析，提高天文学家光

谱处理的效率。

我们会根据需求进一步完善该工具功能，借鉴国际上优秀的光谱可视化软件优势，最终使其成为处理 LAMOST、SDSS 等光谱数据的首选工具。

参考文献

[1]施建荣.LAMOST 望远镜[J].科学通报,2016,61(12):1330-1335.

[2]毛晓艳,张博,叶中付.基于加权滤波的低信噪比 LAMOST 光纤光谱信号降噪[J].天文研究与技术,2015,12(4) : 447-454.

[3] Osuna P, Barbarisi I, SalgadoJ, et al. VOSpec: a tool for handling virtual observatory compliant spectra [C] // Astronomical Society of the Pacific Conference Series. 2005: 198202.

[4] Busko I. Specview: a Java tool for spectral visualization and model fiting [C] //Proceedings of the SPIE. 2002: 410-418.

[5] Škoda P, PerP W, Neves M C, et al. Spectroscopic analysis in the virtual observator environment with SPLAT-VO [J]. Astronomy & Computing, 2014, 7: 108-120.

[6] Lebouteiller V, Barry D J, Spoon H W W, et al. CASSIS: the cornell atlas of spitzer / infrared spectrograph sources []. Astrophysical Journal Supplement, 2011，196(1): 849-856.

[7] Yuan Hailong, Zhang Haotong, Zhang Yanxia, et al. ASERA: a spectrum eye recognition assistant for quasar spectra [J]. Astronomy & Computing, 2013, 3-4: 65-69.

[8]雷国洪,徐洋,牛晨辉,田海俊,张彦霞,崔辰州,赵永恒.专家识谱平台的设计与实现[J].天文研究与技术,2018,15(02):216-224.

[9]王文成,李健,王瑞兰,吴小进,孙学岩.基于 Matlab GUI 的数字图像处理仿真平台设计与开发[J].实验技术与管理,2019,36(02):141-144.

[10]彭广民,陈婷.基于 Matlab 小波去噪的研究方法[J].测绘与空间地理息,2016,39(07):24-26.

[11]李天超,邓李才,赵刚.漩涡星系 M31 和椭圆星系 M32CaⅡ三重线的等值宽度[J].光谱学与光谱分析,2000(04):468-470.

[12]刘蓉,段福庆,刘三阳,吴福朝.基于知识的红移测量和谱线证认方法[J].电子与信息学报，2006(01):76-79.

[13]屠良平,罗阿理,姜斌,韦鹏,赵永恒,刘蓉.一种新的活动星系光谱红移自动测量方法[J].光谱学与光谱分析,2012,32(10):2858-2862.